



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/021,368

DATE: 02/19/2002

TIME: 11:52:38

Input Set : N:\Crf3\RULE60\10021368.txt
 Output Set: N:\CRF3\02192002\J021368.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Band, Vimla

8 (ii) TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
9 MOLECULES AND METHODS

11 (iii) NUMBER OF SEQUENCES: 11

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: Fish & Richardson P.C.

15 (B) STREET: 225 Franklin Street

16 (C) CITY: Boston

17 (D) STATE: MA

18 (E) COUNTRY: USA

19 (F) ZIP: 02110-2804

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Floppy disk

23 (B) COMPUTER: IBM PC compatible

24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/10/021,368

C--> 29 (B) FILING DATE: 12-Dec-2001

34 (C) CLASSIFICATION:

31 (vii) PRIOR APPLICATION DATA:

32 (A) APPLICATION NUMBER: 09/201,038

33 (B) FILING DATE:

36 (viii) ATTORNEY/AGENT INFORMATION:

37 (A) NAME: Clark, Paul T.

38 (B) REGISTRATION NUMBER: 30,162

39 (C) REFERENCE/DOCKET NUMBER: 00398/100002

41 (ix) TELECOMMUNICATION INFORMATION:

42 (A) TELEPHONE: 617/542-5070

43 (B) TELEFAX: 617/542-8906

44 (C) TELEX: 200154

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 276 amino acids

51 (B) TYPE: amino acid

52 (C) STRANDEDNESS: not relevant

53 (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: protein

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

62 Met Arg Ala Pro His Leu His Leu Ser Ala Ala Ser Gly Ala Arg Ala

63 1 5 10 15

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/021,368

DATE: 02/19/2002
TIME: 11:52:38

Input Set : N:\CrF3\RULE60\10021368.txt
Output Set: N:\CRF3\02192002\J021368.raw

65 Leu Ala Lys Leu Leu Pro Leu Leu Met Ala Gln Leu Trp Ala Ala Glu
66 20 25 30
69 Ala Ala Leu Leu Pro Gln Asn Asp Thr Arg Leu Asp Pro Glu Ala Tyr
70 35 40 45
72 Gly Ala Pro Cys Ala Arg Gly Ser Gln Pro Trp Gln Val Ser Leu Phe
73 50 55 60
75 Asn Gly Leu Ser Phe His Cys Ala Gly Val Leu Val Asp Gln Ser Trp
76 65 70 75 80
78 Val Leu Thr Ala Ala His Cys Gly Asn Lys Pro Leu Trp Ala Arg Val
79 85 90 95
81 Gly Asp Asp His Leu Leu Leu Gln Gly Glu Gln Leu Arg Arg Thr
82 100 105 110
84 Thr Arg Ser Val Val His Pro Lys Tyr His Gln Gly Ser Gly Pro Ile
85 115 120 125
87 Leu Pro Arg Arg Thr Asp Glu His Asp Leu Met Leu Leu Lys Leu Ala
88 130 135 140
90 Arg Pro Val Val Pro Gly Pro Arg Val Arg Ala Leu Gln Leu Pro Tyr
91 145 150 155 160
93 Arg Cys Ala Gln Pro Gly Asp Gln Cys Gln Val Ala Gly Trp Gly Thr
94 165 170 175
96 Thr Ala Ala Arg Arg Val Lys Tyr Asn Lys Gly Leu Thr Cys Ser Ser
97 180 185 190
99 Ile Thr Ile Leu Ser Pro Lys Glu Cys Glu Val Phe Tyr Pro Gly Val
100 195 200 205
102 Val Thr Asn Asn Met Ile Cys Ala Gly Leu Asp Arg Gly Gln Asp Pro
103 210 215 220
105 Cys Gln Ser Asp Ser Gly Gly Pro Leu Val Cys Asp Glu Thr Leu Gln
106 225 230 235 240
108 Gly Ile Leu Ser Trp Gly Val Tyr Pro Cys Gly Ser Ala Gln His Pro
109 245 250 255
111 Ala Val Tyr Thr Gln Ile Cys Lys Tyr Met Ser Trp Ile Asn Lys Val
112 260 265 270
114 Ile Arg Ser Asn
115 275

117 (2) INFORMATION FOR SEQ ID NO: 2:

119 (i) SEQUENCE CHARACTERISTICS:

120 (A) LENGTH: 1454 base pairs
121 (B) TYPE: nucleic acid
122 (C) STRANDEDNESS: single
123 (D) TOPOLOGY: linear

125 (ii) MOLECULE TYPE: cDNA

130 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

132 ACCAGCGGCA GACCACAGGC AGGGCAGAGG CACGTCTGGG TCCCCTCCCT CCTTCCTATC	60
134 GCGGACTCCC AGATCCTGGC CATGAGAGCT CCGCACCTCC ACCTCTCGC CGCCTCTGGC	120
136 GCGCGGGCTC TGGCGAAGCT GCTGCCGCTG CTGATGGCGC AACTCTGGGC CGCAGAGGCG	180
138 GCGCTGCTCC CCCAAACGA CACCGCGCTTG GACCCCGAAG CCTATGGCGC CGCGTGCCTG	240
140 CGCGGCTCGC AGCCCTGGCA GGTCTCGCTC TTCAACGGCC TCTCGTTCCA CTGCGCGGGT	300
142 GTCTCTGGTGG ACCAGAGTTG GGTGCTGACG GCCGCGCACT GCGGAAACAA GCCACTGTGG	360
144 GCTCGAGTAG GGGATGATCA CCTGCTGCTT CTTCAAGGGCG AGCAGCTCCG CCGGACGACT	420

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/021,368

DATE: 02/19/2002
TIME: 11:52:38

Input Set : N:\Crf3\RULE60\10021368.txt
Output Set: N:\CRF3\02192002\J021368.raw

146	CGCTCTGTTG	TCCATCCCAA	GTACCACCAAG	GGCTCAGGCC	CCATCCTGCC	AAGGCGAACG	480
148	GATGAGCACG	ATCTCATGTT	GCTAAAGCTG	GCCAGGCCG	TAGTGCCGGG	GCCCCGGCTC	540
150	CGGGCCCTGC	AGCTTCCCTA	CCGCTGTGCT	CAGCCGGAG	ACCAGTGCCA	GGTTGCTGGC	600
152	TGGGGCACCA	CGGCCGCCG	GAGAGTGAAG	TACAACAAGG	GGCTGACCTG	CTCCAGCATE	660
154	ACTATCCTGA	GCCCTAAAGA	GTGTGAGGTG	TTCTACCCCTG	GCGTGGTCAC	CAACAAACATG	720
156	ATATGTGCTG	GACTGGACCG	GGGCCAGGAC	CTTGCCAGA	GTGACTCTGG	AGGCCCOCTG	780
158	GTCTGTGACG	AGACCCCTCCA	AGGCATCCTC	TCGTGGGTG	TTTACCCCTG	TGGCTCTGCC	840
160	CAGCATCCAG	CTGTCTACAC	CCAGATCTGC	AAATACATGT	CCTGGATCAA	TAAAGTCATA	900
162	CGCTCCAAC	GATCCAGATG	CTACGCTCCA	GCTGATCCAG	ATGTTATGCT	CCTGCTGATC	960
164	CAGATGCCCA	GAGGCTCCAT	CGTCCATCCT	CITCCTCCCC	AGTCGGCTGA	ACTCTCCCT	1020
166	TGTCTGCACT	GTTCAAACCT	CTGCCGCCCT	CCACACCTCT	AAACATCTCC	CCTCTCACCT	1080
168	CATTCCCCCA	CCTATCCCCA	TTCTCTGCT	GTACTGAAGC	TGAAATGCAG	GAAGTGGTGG	1140
170	CAAAGGTTA	TTCCAGAGAA	GCCAGGAAGC	CGGTCACTAC	CCAGCCCTCG	AGAGCAGTTA	1200
172	CTGGGGTCAC	CCAACCTGAC	TTCTCTGCC	ACTCCCCGCT	GTGTGACTTI	GGGCAAGCCA	1260
174	AGTCCCCCT	CTGAACCTCA	GTTCCTCAT	CTGCAAAATG	GGAACAATGA	CGTGCCTACC	1320
176	TCITAGACAT	GTGTGAGGA	GACTATGATA	TAACATGTGT	ATGTAATCT	TCATGTGATT	1380
178	GTCATGTAAG	GCTTAACACA	GTGGGTGGTG	AGTTCTGACT	AAAGGTTACC	TGTTGTCGTG	1440
180	AAAAAAAAAA	AAAA					1454

182 (2) INFORMATION FOR SEQ ID NO: 3:

184 (i) SEQUENCE CHARACTERISTICS:
185 (A) LENGTH: 19 base pairs
186 (B) TYPE: nucleic acid
187 (C) STRANDEDNESS: single
188 (D) TOPOLOGY: linear

189 (ii) MOLECULE TYPE: cDNA

190 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

191 CGCGAGATT AGGTGACAC

19

192 (2) INFORMATION FOR SEQ ID NO: 4:

193 (i) SEQUENCE CHARACTERISTICS:
194 (A) LENGTH: 19 base pairs
195 (B) TYPE: nucleic acid
196 (C) STRANDEDNESS: single
197 (D) TOPOLOGY: linear

198 (ii) MOLECULE TYPE: cDNA

199 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

200 GGCCTCTAAC ACGACTCAC

19

201 (2) INFORMATION FOR SEQ ID NO: 5:

202 (i) SEQUENCE CHARACTERISTICS:
203 (A) LENGTH: 17 base pairs
204 (B) TYPE: nucleic acid
205 (C) STRANDEDNESS: single
206 (D) TOPOLOGY: linear

207 (ii) MOLECULE TYPE: cDNA

208 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

209 TACCACTACA ATGGATG

17

210 (2) INFORMATION FOR SEQ ID NO: 6:

211 (i) SEQUENCE CHARACTERISTICS:
212 (A) LENGTH: 20 base pairs
213 (B) TYPE: nucleic acid

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/021,368

DATE: 02/19/2002
TIME: 11:52:38

Input Set : N:\Crf3\RULE60\10021368.txt
Output Set: N:\CRF3\02192002\J021368.raw

238 (C) STRANDEDNESS: single
239 (D) TOPOLOGY: linear
241 (ii) MOLECULE TYPE: cDNA
246 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6: 20
248 GTTGAAGTGA ACTTGCGGGC
250 (2) INFORMATION FOR SEQ ID NO: 7:
252 (i) SEQUENCE CHARACTERISTICS:
253 (A) LENGTH: 281 amino acids
254 (B) TYPE: amino acid
255 (C) STRANDEDNESS: not relevant
256 (D) TOPOLOGY: linear
258 (ii) MOLECULE TYPE: protein
263 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
265 Met Ser Ala Leu Leu Ile Leu Ala Leu Gly Ala Ala Val Ala Phe
266 1 5 10 15
266 Pro Val Asp Asp Asp Asp Lys Ile Val Gly Gly Tyr Met Met Thr Arg
268 20 25 30
269 Tyr Ala Arg Thr Cys Arg Glu Ser Ser Val Pro Tyr Gln Val Ser Leu
271 35 40 45
272 Asn Ala Gly Tyr His Phe Cys Gly Gly Ser Leu Ile Asn Asp Gln Trp
274 50 55 60
275 Val Val Ser Ala Ala His Cys Tyr Lys Tyr Arg Ile Gln Val Arg Leu
277 65 70 75 80
278 Gly Glu His Asn Met Met Thr Arg Tyr Ala Arg Ile Asn Val Leu Glu
280 85 90 95
281 Gly Asn Glu Gln Phe Val Asp Ser Ala Lys Ile Ile Arg His Pro Asn
283 100 105 110
284 Tyr Asn Ser Trp Thr Leu Asp Asn Asp Ile Met Leu Ile Lys Leu Ala
286 115 120 125
287 Ser Pro Val Thr Leu Met Met Thr Arg Tyr Ala Arg Asn Ala Arg Val
289 130 135 140
290 Ala Ser Val Pro Leu Pro Ser Ser Cys Ala Pro Ala Gly Thr Gln Cys
292 145 150 155 160
293 Leu Ile Ser Gly Trp Gly Asn Thr Leu Ser Asn Gly Val Asn Asn Pro
295 165 170 175
296 Asp Leu Leu Gln Cys Val Asp Ala Pro Val Leu Pro Gln Ala Met Met
298 180 185 190
299 Thr Arg Tyr Ala Arg Asp Cys Glu Ala Ser Tyr Pro Gly Asp Ile Thr
301 195 200 205
302 Asn Asn Met Ile Cys Val Gly Phe Leu Glu Gly Gly Lys Asp Ser Cys
304 210 215 220
305 Gln Gly Asp Ser Gly Gly Pro Val Val Cys Asn Gly Glu Leu Gln Gly
307 225 230 235 240
308 Ile Val Ser Trp Gly Tyr Met Met Thr Arg Tyr Ala Arg Gly Cys Ala
310 245 250 255
311 Gln Pro Asp Ala Pro Gly Val Tyr Thr Lys Val Cys Asn Tyr Val Asp
313 260 265 270
314 Trp Ile Gln Asn Thr Ile Ala Asp Asn
316 275 280
317

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/021,368

DATE: 02/19/2002
TIME: 11:52:38

Input Set : N:\Crf3\RULE60\10021368.txt
Output Set: N:\CRF3\02192002\J021368.raw

319 (2) INFORMATION FOR SEQ ID NO: 8:

321 (i) SEQUENCE CHARACTERISTICS:

322 (A) LENGTH: 299 amino acids

323 (B) TYPE: amino acid

324 (C) STRANDEDNESS: not relevant

325 (D) TOPOLOGY: linear

327 (ii) MOLECULE TYPE: protein

332 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Leu Glu Leu His Pro Leu Leu Gly Gly Arg Thr Trp Arg Ala Ala Arg
1 5 10 15
Asp Ala Asp Gly Cys Glu Ala Leu Gly Thr Val Ala Val Pro Phe Asp
20 25 30
Asp Asp Asp Lys Ile Val Gly Gly Tyr His Ser Thr Arg Tyr Ile Val
35 40 45
Asx Thr Cys Glu Asn Ser Leu Pro Tyr Gln Val Ser Leu Asn Ser Gly
50 55 60
Ser His Phe Cys Gly Gly Ser Leu Ile Ser Glu Gln Trp Val Val Ser
65 70 75 80
Ala Ala His Cys Tyr Lys Thr Arg Ile Gln Val Arg Leu Gly Glu His
85 90 95
Asn His Ser Thr Arg Tyr Ile Val Asx Ile Lys Val Leu Glu Gly Asn
100 105 110
Glu Gln Phe Ile Asn Ala Ala Lys Ile Ile Arg His Pro Lys Tyr Asn
115 120 125
Arg Asp Thr Leu Asp Asn Asp Ile Met Leu Ile Lys Leu Ser Ser Pro
130 135 140
Ala Val Ile His Ser Thr Arg Tyr Ile Val Asx Asn Ala Arg Val Ser
145 150 155 160
Thr Ile Ser Leu Pro Thr Ala Pro Pro Ala Ala Gly Thr Glu Cys Leu
165 170 175
Ile Ser Gly Trp Gly Asn Thr Leu Ser Phe Gly Ala Asp Tyr Pro Asp
180 185 190
Glu Leu Lys Cys Leu Asp Ala Pro Val Leu Thr Gln Ala His Ser Thr
195 200 205
Arg Tyr Ile Val Asx Glu Cys Lys Ala Ser Tyr Pro Gly Lys Ile Thr
210 215 220
Asn Ser Met Phe Cys Val Gly Phe Leu Glu Gly Gly Lys Asp Ser Cys
225 230 235 240
Gln Arg Asp Ser Gly Gly Pro Val Val Cys Asn Gly Gln Leu Gln Gly
245 250 255
Val Val Ser Trp Gly His His Ser Thr Arg Tyr Ile Val Asx Gly Cys
260 265 270
Ala Trp Lys Asn Arg Pro Gly Val Tyr Thr Lys Val Tyr Asn Tyr Val
275 280 285
Asp Trp Ile Lys Asp Thr Ile Ala Ala Asn Ser
290 295

391 (2) INFORMATION FOR SEQ ID NO: 9:

393 (i) SEQUENCE CHARACTERISTICS:

394 (A) LENGTH: 286 amino acids

VERIFICATION SUMMARY

PATENT APPLICATION: **US/10/021,368**

DATE: 02/19/2002

TIME: 11:52:39

Input Set : **N:\Crf3\RULE60\10021368.txt**
Output Set: **N:\CRF3\02192002\J021368.raw**

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]